



USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)Search: ☐ The ACM Digital Library ☒ The Guide**THE GUIDE TO COMPUTING LITERATURE** [Feedback](#)**Rate of change and other metrics: a live study of the world wide web**

Full text Pdf (251 KB)

Source **Proceedings of the USENIX Symposium on Internet Technologies and Systems on USENIX Symposium on Internet Technologies and Systems** [table of contents](#)

Monterey, California

Pages: 14 - 14

Year of Publication: 1997

Authors **Fred Douglass** AT&T Labs-Research
Anja Feldmann AT&T Labs-Research
Balachander Krishnamurthy AT&T Labs-Research
Jeffrey Mogul Digital Equipment Corporation, Western Research Laboratory

Publisher: USENIX Association Berkeley, CA, USA

Additional Information: [abstract](#) [references](#) [cited by](#) [index terms](#) [collaborative colleagues](#)

Tools and Actions: [Review this Article](#)
[Save this Article to a Binder](#) Display Formats: [BibTex](#) [EndNote](#) [ACM Ref](#)

↑ ABSTRACT

Caching in the World Wide Web is based on two critical assumptions: that a significant fraction of requests reaccess resources that have already been retrieved; and that those resources do not change between accesses.

We tested the validity of these assumptions, and their dependence on characteristics of Web resources, including access rate, age at time of reference, content type, resource size, and Internet top-level domain. We also measured the rate at which resources change, and the prevalence of duplicate copies in the Web.


We quantified the potential benefit of a shared proxy-caching server in a large environment by using traces that were collected at the Internet connection points for two large corporations, representing significant numbers of references. Only 22% of the resources referenced in the traces we analyzed were accessed more than once, but about half of the references were to those multiply-referenced resources. Of this half, 13% were to a resource that had been modified since the previous traced reference to it.

We found that the content type and rate of access have a strong influence on these metrics, the domain has a moderate influence, and size has little effect. In addition, we studied other aspects of the rate of change, including semantic differences such as the insertion or deletion of anchors, phone numbers, and email addresses.



↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has





opted to expose the complete List rather than only correct and linked references.

- 1 [1] Martin F. Arlitt and Carey L. Williamson. Web server workload characterization: The search for invariants (extended version). Technical Report DISCUS Working Paper 96-3, Dept. of Computer Science, University of Saskatchewan, March 1996. Available as <ftp://ftp.cs.usask.ca/pub/discus/paper.96.3.ps.Z>.
- 2 [2] Gaurav Banga, Fred Dougliis, and Michael Rabinovich. Optimistic deltas for WWW latency reduction. In *Proceedings of 1997 USENIX Technical Conference*, pages 289-303, Anaheim, CA, January 1997. Also available as <http://www.research.att.com/~dougliis/papers/optdel.ps.gz>.
- 3 Azer Bestavros, Speculative Data Dissemination and Service to Reduce Server Load, Network Traffic and Service Time in Distributed Information Systems, Proceedings of the Twelfth International Conference on Data Engineering, p.180-187, February 26-March 01, 1996
- 4 Tim Bray, Measuring the Web, Proceedings of the fifth international World Wide Web conference on Computer networks and ISDN systems, p.993-1005, May 1996, Paris, France
- 5 Andrei Z. Broder , Steven C. Glassman , Mark S. Manasse , Geoffrey Zweig, Syntactic clustering of the Web, Selected papers from the sixth international conference on World Wide Web, p.1157-1166, September 1997, Santa Clara, California, United States
- 6 Carlos Cunha , Azer Bestavros , Mark Crovella, Characteristics of WWW Client-based Traces, Boston University, Boston, MA, 1995
- 7 Fred Dougliis , Thomas Ball , Yih-farn Chen , Eleftherios Koutsofios, The AT&T Internet Difference Engine: Tracking and viewing changes on the web, World Wide Web, v.1 n.1, p.27-44, 1998 [doi>10.1023/A:1019243126596]
- 8 [8] Fred Dougliis, Anja Feldmann, Balachander Krishnamurthy, and Jeffrey Mogul. Rate of change and other metrics: a live study of the World Wide Web. Technical Report #97.24.2, AT&T Labs-Research, Florham Park, NJ, December 1997. Available as <http://www.research.att.com/library/trs-/TRs/97/97.24/97.24.2.body.ps>.
- 9 R. Fielding , J. Gettys , J. Mogul , H. Frystyk , T. Berners-Lee, Hypertext Transfer Protocol -- HTTP/1.1, RFC Editor, 1997
- 10 [10] Steven D. Gribble and Eric A. Brewer. System design issues for internet middleware services: Deductions from a large client trace. In *Proceedings of the Symposium on Internetworking Systems and Technologies*. USENIX, December 1997. To appear.
- 11 [11] James Gwertzman and Margo Seltzer. World-Wide Web cache consistency. In *Proceedings of 1996 USENIX Technical Conference*, pages 141-151, San Diego, CA, January 1996. Also available as <http://www.eecs.harvard.edu/~vino/web/usenix.196/>.
-  12 Barron C. Housel , David B. Lindquist, WebExpress: a system for optimizing Web browsing in a wireless environment, Proceedings of the 2nd annual international conference on Mobile computing and networking, p.108-116, November 1996, Rye, New York, United States [doi>10.1145/236387.236416]
- 13 [13] Inktomi. <http://inktom.berkeley.edu>, January 1997.
- 14 [14] Guy Jacobson, Balachander Krishnamurthy, and Divesh Srivastava. Grink: To grok and link. Technical Memorandum, AT&T Labs-Research, July 1996.
- 15 [15] Thomas M. Kroeger, Darrell D. E. Long, and Jeffrey C. Mogul. Exploring the bounds of web latency reduction from caching and prefetching. In *Proceedings of the Symposium on*

Internetworking Systems and Technologies . USENIX, December 1997. To appear.
Available as <http://WWW.cse.ucsc.edu/~tmk/ideal.ps>.

- 16  Jeffrey C. Mogul , Fred Douglass , Anja Feldmann , Balachander Krishnamurthy, Potential benefits of delta encoding and data compression for HTTP, Proceedings of the ACM SIGCOMM '97 conference on Applications, technologies, architectures, and protocols for computer communication, p.181-194, September 14-18, 1997, Cannes, France
- 17 [17] Url-minder. <http://www.netmind.com/URL-minder/URL-minder.html>, December 1996.
- 18 [18] Opentext. <http://www.opentext.com>, 1997.
- 19 [19] Arthur van Hoff, John Giannandrea, Mark Hapner, Steve Carter, and Milo Medin. The http distribution and replication protocol. W3C Note, available as <http://www.w3.org/TR/NOTE-drp-19970825.html>, August 1997.
- 20 [20] Charles L. Viles and James C. French. Availability and latency of World Wide Web information servers. *Computing Systems*, 8(1):61-91, Winter 1995.
- 21  Marc Abrams , Charles R. Standridge , Ghaleb Abdulla , Edward A. Fox , Stephen Williams, Removal policies in network caches for World-Wide Web documents, Conference proceedings on Applications, technologies, architectures, and protocols for computer communications, p.293-305, August 28-30, 1996, Palo Alto, California, United States
- 22 Allison Woodruff , Paul M. Aoki , Eric Brewer , Paul Gauthier, An investigation of documents from the World Wide Web, Proceedings of the fifth international World Wide Web conference on Computer networks and ISDN systems, p.963-980, May 1996, Paris, France

↑ CITED BY 6

-  Herman Chung-Hwa Rao , Yih-Farn Chen , Ming-Feng Chen, A proxy-based personal web archiving service, ACM SIGOPS Operating Systems Review, v.35 n.1, p.61-72, January 1, 2001
-  Anirban Dasgupta , Arpita Ghosh , Ravi Kumar , Christopher Olston , Sandeep Pandey , Andrew Tomkins, The discoverability of the web, Proceedings of the 16th international conference on World Wide Web, May 08-12, 2007, Banff, Alberta, Canada
-  Ziv Bar-Yossef , Idit Keidar , Uri Schonfeld, Do not crawl in the dust: different urls with similar text, Proceedings of the 16th international conference on World Wide Web, May 08-12, 2007, Banff, Alberta, Canada
- Timothy K. Shih , Shi-Kuo Chang , Jeffrey Tsai , Jianhua Ma , Runhe Huang, Supporting Well-Engineered Web Documentation Development – a Multimedia Software Engineering Approach toward Virtual University Courseware Designs, Annals of Software Engineering, v.12 n.1, p.139-165, December 2001
-  Frank McCown , Norou Diawara , Michael L. Nelson, Factors affecting website reconstruction from the web infrastructure, Proceedings of the 2007 conference on Digital libraries, June 18-23, 2007, Vancouver, BC, Canada
- Abdullah Balamash , Marwan Krunz , Philippe Nain, Performance analysis of a client-side caching/prefetching system for Web traffic, Computer Networks: The International Journal of Computer and Telecommunications Networking, v.51 n.13, p.3673-3692, September, 2007

↑ INDEX TERMS

Primary Classification:

C. Computer Systems Organization↳ **C.2 COMPUTER-COMMUNICATION NETWORKS**↳ **C.2.4 Distributed Systems****Additional Classification:****H. Information Systems**↳ **H.3 INFORMATION STORAGE AND RETRIEVAL**↳ **H.3.4 Systems and Software**↳ **Subjects:** Performance evaluation (efficiency and effectiveness)↳ **Nouns:** World Wide Web (WWW)**General Terms:**

Measurement, Performance

↑ **Collaborative Colleagues:**

<u>Frederick Douglass</u>	<u>Thomas Jaudon Ball</u>	<u>A van Hoff</u>	<u>Ling Liu</u>	<u>Bill N Schilit</u>
<u>Thomas Ball</u>	<u>Jim Isaak</u>	<u>Zhuoqing Morley</u>	<u>James Paul Sienicki</u>	<u>Oliver Spatscheck</u>
<u>Gaurav</u>	<u>Arun K Iyengar</u>	<u>Mao</u>	<u>Brian Marsh</u>	<u>Madhusudhan Talluri</u>
<u>Banga</u>	<u>Sonia Jain</u>	<u>Dejan S</u>	<u>David Tao</u>	<u>Joshua A Tauber</u>
<u>Michael</u>	<u>Frans Kaashoek</u>	<u>Milo&jacute;icic</u>	<u>Joshua A Tauber</u>	<u>William Harold Tetzlaff</u>
<u>Branson</u>	<u>Chuck Kalmanek</u>	<u>Dejan S Milojicic</u>	<u>John Michael Tracey</u>	<u>John A Trotter</u>
<u>Ramon</u>	<u>Pradnya Karbhari</u>	<u>Dejan</u>	<u>John A Trotter</u>	<u>Jia Wang</u>
<u>Caceres</u>	<u>Tom Killian</u>	<u>Milojicicacute</u>	<u>Brent Ballinger Welch</u>	<u>Richard G Wheeler</u>
<u>Stephen Joel</u>	<u>John Klensin</u>	<u>Jeffrey C Mogul</u>	<u>Cathy H Xia</u>	<u>Zhen Xiao</u>
<u>Chapin</u>	<u>Eleftherios E</u>	<u>Bruce R Montague</u>	<u>Fan Ye</u>	<u>Jian Yin</u>
<u>Yih Farn</u>	<u>Koutsofios</u>	<u>Michael N Nelson</u>	<u>Songnian Zhou</u>	
<u>Robin Chen</u>	<u>Balachander</u>	<u>John Kenneth</u>		
<u>Andrew R</u>	<u>Krishnamurthy</u>	<u>Ousterhout</u>		
<u>Cherenson</u>	<u>David M Kristol</u>	<u>Yves Paindaveine</u>		
<u>Charles D</u>	<u>Paul Krzyzanowski</u>	<u>John Palmer</u>		
<u>Cranor</u>	<u>Purushottam</u>	<u>Michael Rabinovich</u>		
<u>Brian</u>	<u>Kulkarni</u>	<u>Prabhakar</u>		
<u>Douglas</u>	<u>Dinesh</u>	<u>Raghavan</u>		
<u>Davison</u>	<u>Chandrakant</u>	<u>Lakshmish</u>		
<u>Allan Ellis</u>	<u>Kulkarni</u>	<u>Macheeri</u>		
<u>Anja</u>	<u>Ravindra</u>	<u>Ramaswamy</u>		
<u>Feldmann</u>	<u>Ramachandria</u>	<u>Elizabeth S</u>		
<u>Gideon Glass</u>	<u>Kuramkote</u>	<u>Richards</u>		
<u>Y Goland</u>	<u>Jason LaVoie</u>	<u>Bin Rong</u>		
<u>Tatsuya</u>	<u>Kai Li</u>	<u>Bin Rong</u>		
<u>Hagino</u>				
<u>Antonio Haro</u>				
<u>D Hellerstein</u>				
<u>Kirsten</u>				
<u>Hildrum</u>				
<u>Anja Feldmann</u>	<u>Vinay</u>	<u>Polly Huang</u>	<u>Wolfgang</u>	<u>Shang Hua Hua Teng</u>
	<u>Aggarwal</u>	<u>Youngmi Joo</u>	<u>Muhlbauer</u>	<u>Shang Hua Hua Teng</u>
	<u>Arthur W</u>	<u>Nabil Kahale</u>	<u>David Richard O</u>	<u>Dirk Theune</u>
	<u>Berger</u>	<u>Nils Kammenhuber</u>	<u>Hallaron</u>	<u>R Thiele</u>
	<u>Robert</u>	<u>Ming Yang Yang</u>	<u>Vern Edward</u>	<u>Fred True</u>
	<u>Bradford</u>	<u>Kao</u>	<u>Paxson</u>	<u>Fred True</u>
	<u>Ramon</u>	<u>Ming Hung Kao</u>	<u>Bjorn Poonen</u>	<u>Steve Uhlig</u>

	<u>Caceres</u>	<u>Stefan Kornexl</u>	<u>Michael Rabinovich</u>	<u>Jorg Wallerich</u>
	<u>Edward</u>	<u>Evangelos</u>	<u>Nicholas Frederick</u>	<u>Thomas Edward</u>
	<u>Grady</u>	<u>Kotsovinos</u>	<u>Reingold</u>	<u>Warfel</u>
	<u>Coffman</u>	<u>Balachander</u>	<u>Jennifer Rexford</u>	<u>Gerhard Weikum</u>
	<u>Jon</u>	<u>Krishnamurthy</u>	<u>Vinay J Ribeiro</u>	<u>David J Wetherall</u>
	<u>Crowcroft</u>	<u>Thomas G Kurtz</u>	<u>Matthew Roughan</u>	<u>Ward Whitt</u>
	<u>Roberto De</u>	<u>Thomas Lengauer</u>	<u>Andy Rupp</u>	<u>Arne Wichmann</u>
	<u>Prisco</u>	<u>Carsten Lund</u>	<u>Christian</u>	<u>Walter Willinger</u>
	<u>Christian</u>	<u>Julia Luxemburger</u>	<u>Scheideler</u>	<u>Martina Zitterbart</u>
	<u>Dewes</u>	<u>Olaf Maennel</u>	<u>Harald Schioberg</u>	
	<u>Frederick</u>	<u>Bruce M Maggs</u>	<u>Jiri Sgall</u>	
	<u>Douglis</u>	<u>Michael Mai</u>	<u>Jiri Sgall</u>	
	<u>Holger</u>	<u>Zhuoqing Morley</u>	<u>Robin Sommer</u>	
	<u>Dreger</u>	<u>Mao</u>	<u>Robin Sommer</u>	
	<u>Anna C</u>	<u>Jeffrey C Mogul</u>	<u>Thomas M Stricker</u>	
	<u>Gilbert</u>		<u>Ravi Sundaram</u>	
	<u>Gideon Glass</u>			
	<u>Y Goland</u>			
	<u>Albert</u>			
	<u>Gordon</u>			
	<u>Greenberg</u>			
	<u>Thomas</u>			
	<u>Gross</u>			
	<u>D Hellerstein</u>			
	<u>A van Hoff</u>			
	<u>Polly Huang</u>			
<u>Balachander</u>	<u>Aditya Akella</u>	<u>Emden R Gansner</u>	<u>Harsha V</u>	<u>John J Snyder</u>
<u>Krishnamurthy:</u>	<u>Mark Allman</u>	<u>Y Goland</u>	<u>Madhyastha</u>	<u>Oliver Spatscheck</u>
	<u>Martin F</u>	<u>Matthias</u>	<u>Delfina Malandrino</u>	<u>Divesh Srivastava</u>
	<u>Arlitt</u>	<u>Grossglauser</u>	<u>Jeffrey C Mogul</u>	<u>Dan Suciu</u>
	<u>Paul Robert</u>	<u>Robert E Gruber</u>	<u>Euthimios Panagos</u>	<u>Dave W Thomas</u>
	<u>Barford</u>	<u>Robert E Gruber</u>	<u>Jeffrey Pang</u>	<u>Suresh</u>
	<u>Naser Saleh</u>	<u>D Hellerstein</u>	<u>Raghavan</u>	<u>Venkatasubramanian</u>
	<u>Barghouti</u>	<u>A van Hoff</u>	<u>Pratiwadi</u>	<u>Kashi Venkatesh</u>
	<u>David</u>	<u>Paola Inverardi</u>	<u>Michael Rabinovich</u>	<u>Vishwanath</u>
	<u>Belanger</u>	<u>Guy Joseph</u>	<u>Pratap</u>	<u>Kiem Phong Phong Vo</u>
	<u>Yih Farn</u>	<u>Jacobson</u>	<u>Ramamurthy</u>	<u>Jorg Wallerich</u>
	<u>Robin Chen</u>	<u>Jaeyeon Jung</u>	<u>Jennifer Rexford</u>	<u>Jia Wang</u>
	<u>Yan Chen</u>	<u>Dina Katabi</u>	<u>Lynn M Rivera</u>	<u>Walter Willinger</u>
	<u>Edith Cohen</u>	<u>Sachin Katti</u>	<u>David Samuel</u>	<u>Craig Ellis Wills</u>
	<u>Charles D</u>	<u>John T Korb</u>	<u>Rosenblum</u>	<u>Craig E Wills</u>
	<u>Cranor</u>	<u>Eleftherios E</u>	<u>Vyas Sekar</u>	<u>Yinglian Xie</u>
	<u>Frederick</u>	<u>Koutsofios</u>	<u>Subhabrata Sen</u>	<u>Daniel Yankelevich</u>
	<u>Douglis</u>	<u>David M Kristol</u>	<u>Srinivasan Seshan</u>	<u>Yin Zhang</u>
	<u>Holger</u>	<u>Richard Liston</u>	<u>Anees A Shaikh</u>	
	<u>Dreger</u>		<u>Mohammed Shakir</u>	
	<u>Nick G</u>			
	<u>Duffield</u>			
	<u>Anja</u>			
	<u>Feldmann</u>			
	<u>Richard N</u>			
	<u>Ferri</u>			
	<u>Glenn</u>			
	<u>Stephen</u>			
	<u>Fowler</u>			
<u>Jeffrey C Mogul:</u>	<u>Vinay</u>	<u>Anja Feldmann</u>	<u>P Leach</u>	<u>KK K Ramakrishnan</u>
	<u>Aggarwal</u>	<u>R Fielding</u>	<u>Darrell D E Long</u>	<u>Patrick Reynolds</u>

Marcos	H Frystyk	David Ellis Lowell	Patrick Reynolds
Kawazoe	J Gettys	Honghui Lu	Allyn Romanow
Aguilera	Y Goland	Olaf Maennel	Yasushi Saito
Martin F	D Hellerstein	L Masinter	Mehul A Shah
Arlitt	A van Hoff	Robert Nelson	Amitabh Srivastava
Martin Arlitt	Bryan Hopkins	Mayo	Dinesh Subhraveti
Gaurav	Terence Kelly	J McCann	Amin M Vahdat
Banga	Christopher Angel	K McCloghrie	Ben Verghese
Joel F	Kent	Greg Minshall	Janet Lynn Wiener
Bartlett	Charles Edwin	Justin Moore	Janet L Wiener
T Berners	Killian	Athicha	
Lee	Bjorn Knutsson	Muthitacharoen	
Brian Nathan	Balachander	Venkata Narayana	
Bershad	Krishnamurthy	Padmanabhan	
David Reeves	David M Kristol	C Partridge	
Boggs	Thomas M Kroeger	J Postel	
Anita Borg			
Lawrence			
Sivert			
Brakmo			
Yee Man			
Chan			
S E Deering			
Frederick			
Douglis			
Peter			
Druschel			

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2008 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)